Biobanking for Metabolomics and Lipidomics in Precision Medicine

Kirwan, Jennifer A.

2019-07

Kirwan , J A , Kaddurah-Daouk , R , Mitchell , T , Pischon , T , Schmidt , M A & Velagapudi , V 2019 , ' Biobanking for Metabolomics and Lipidomics in Precision Medicine ' , Clinical Chemistry (Washington, DC) , vol. 65 , no. 7 , pp. 827-832 . https://doi.org/10.1373/clinchem.2018.298620

http://hdl.handle.net/10138/304523
https://doi.org/10.1373/clinchem.2018.298620

publishedVersion

Clinical Chemistry 65:7
827–832 (2019)

Q&A

Biobanking for Metabolomics and Lipidomics in Precision Medicine

Moderator: Jennifer A. Kirwan[1,2*]
Experts: Rima Kaddurah-Daouk,[3] Todd Mitchell,[4,5] Tobias Pischon,[6,7,8] Michael A. Schmidt,[9,10] and Vidya Velagapudi[11]

Metabolomics and lipidomics play an important role in precision medicine. Global or targeted analyses of metabolites and lipids provide a specific metabolic phenotype, a "metabotype," of an individual, which can provide important information on their propensity to develop certain diseases or their likely reaction to nutrition or pharmaceuticals. Metabolomics is increasingly being used in research to look at the etiology, prognosis, and outcomes of disease or provide more detailed information on normal human biochemistry. By combining metabolomics data sets with clinical, physiological, or other omics data sets, systems biology approaches can be used to provide context to the observed changes. For human populations in particular, the huge variation in genetics, lifestyle, and other extrinsic factors means that epidemiologic cohorts are often required to reach sufficient statistical power to draw meaningful conclusions. In addition, when disease entities are being studied, it can take years for enough representative samples to be collected. For these reasons, many metabolomics studies are reliant on samples collected and stored in either "official" biobanks (i.e., sample access available to other researchers on application) or privately stored sample repositories.

Biobanks have grown and been increasingly professionalized in the past 20 to 30 years. Many biobanks were originally started for large-scale genetic analyses, and not all repositories have collected biofluids or tissues specifically with metabolomics analyses in mind. Measurements of individual metabolites can be highly dependent on collection and sample storage conditions. Thus, to reliably answer research questions, we need to have reliable collection and storage of samples in a form that makes them suitable for robust metabolomics analyses. To do otherwise invites the introduction of large technical variations in the data set and increases the risk of missing important metabolites that are contributing to the observed phenotypes or incorrectly attributing a metabolite change owing to technical error to a biological cause. Much is still unknown about the optimum way to collect and store biological samples for metabolomics analyses, especially for the less-well-studied tissues and biofluids such as saliva or feces. Here we consult metabolomics, lipids, and biobanking experts to discuss how collection and storage at biobanks can improve metabolomics and lipidomics studies.

What do you think are the most important preanalytical steps in collecting samples for biobanking for metabolomics experiments that contribute to technical variability in the measurement of analytes? Is this the same for all sample types?

Tobias Pischon: Based on my experience as an epidemiologist and investigator in many large-scale cohort studies, as a general rule I think that the more manual work and activity by humans that is involved in the process of collecting and processing samples, and the less these procedures are automated, the higher the likelihood of variability. As such, the collection of samples itself is often the most critical issue. This issue is probably particularly relevant for blood samples, in which the quality of collection heavily depends on the training of technical staff. For example, in the early phase of studies (after their initiation), we still occasionally encounter hemolysis in blood samples, which indicates that training and implementation of blood collection procedures need to be tightly monitored. Similar issues are of course also relevant

[1] Head of the BIH Metabolomics Platform, Max Delbrück Center for Molecular Medicine in the Helmholtz association, Berlin, Germany; [2] Head of the BIH Metabolomics Platform, Berlin Institute of Health, Berlin, Germany; [3] Professor, Psychiatry and Behavioral Sciences, Duke Institute for Brain Sciences, Duke University, Durham, NC; [4] Associate Professor, School of Medicine, University of Wollongong, Wollongong, Australia; [5] Head of the Lipidomics Laboratory, Illawarra Health and Medical Research Institute, Keiraville, Australia; [6] Research Group Head, Molecular Epidemiology Research Group, Max Delbrück Center for Molecular Medicine, Berlin, Germany; [7] Professor, Molecular Epidemiology, Charité Universitätsmedizin, Berlin, Germany; [8] Director, MDC/BIH Biobank, Berlin Institute of Health, Berlin, Germany; [9] Chair, Advanced Pattern Analysis & Countermeasures Group, Boulder, CO; [10] CEO & CSO, Sovaris Aerospace, Boulder, CO; [11] Director and Chair, National Metabolomics Core Facility, Technology Center, Institute for Molecular Medicine FIMM, HiLIFE, University of Helsinki, Finland.
* Address correspondence to this author at: Robert Rössle Strasse 10, Berlin, Germany 13125. E-mail jennifer.kirwan@mdc-berlin.de.
Received February 28, 2019; accepted April 9, 2019.

for other types of samples, such as urine or stool samples, in which the quality heavily depends on how study volunteers follow the instructions for collecting the samples. The processing steps between samples collection and storage (e.g., centrifugation, pipetting) may introduce additional variability. Automation helps to reduce variability. For example, in the German National Cohort health study, pipetting robots have been installed in all recruitment centers, and the process of aliquoting samples has been fully automated.

Vidya Velagapudi: In my opinion, it depends on various factors. For example, a lack of harmonization in the process will increase technical variability, but established quality assurance and control procedures, well-documented standard operating procedures, and trained professionals will improve it. The preanalytical steps vary among different sample types, such as biofluids and tissues, in which each requires specialized handling and storage criteria.

Rima Kaddurah-Daouk: I agree that methods need to be very well laid out and used consistently across sites involved in the collection of samples. Using videos to illustrate methods and writing detailed protocols can help with this, as can training everyone at 1 site before collection starts. Use similar centrifuges if possible. Also important is operating on ice to keep things cold and rapidly processing all samples before transferring to freezers. Record all observations in the processing of samples for each site. And avoid freeze thawing.

Todd Mitchell: If we consider biobanking of tissues, especially if donated after death, I think the key step is ensuring that the postmortem interval is as short as possible and that the samples are snap frozen as quickly as possible. There can be a large increase in oxidative processes and changes to metabolites after circulation stops, so it is important to limit them. Of course, there are other things that are important, and they do vary depending on tissue. For example, in the brain it is important to ensure that the exact region is recorded upon dissection and storage because the cytoarchitecture varies substantially between regions.

Michael A. Schmidt: As a molecular scientist, my work with humans in extreme environments spans variable and difficult field conditions, such as orbital and suborbital spaceflight, military Special Forces operations, athletic teams, high altitude ascent, and standard clinical settings. So, we must be quite attentive to when experimental variance might be introduced. We use a structured approach to address preanalytical variance. First, we break the sources down to variance that may occur at the (*a*) patient, (*b*) physician, and (*c*) laboratory levels, in which each variable class has multiple potential contributors. Across these 3 domains, temperature is among the most crucial elements to control. Second, we look at each individual source of variance in terms of whether we should attempt to control it or merely annotate. This step is necessary as 1 method to reduce false discoveries when multivariate statistics are applied to the resulting high-dimensional data sets.

Although there are many potential sources of experimental variance in each class, a foundational set of variables should be controlled or carefully annotated. For each class, the following warrant consideration:

Patient: Dietary intake, pharmaceuticals, illicit drugs, dietary supplements, botanicals, fasting state, physical activity, occupation, environmental exposure, pregnancy, menstrual status, disease diagnoses, genetic variants, circadian cycles.

Physician: Sampling materials, sampling methods, sampling conditions, sampling hygiene, sampling time of day, number of sampling days, number of sampling locations, number and identity of different phlebotomists, centrifugation, immediate transfer, immediate freezing, temperature of immediate freezing, transport method (e.g., dry ice and amount), transport temperature, days of transport, unusual weather associated with transport (e.g., high temperature), number of freeze–thaw cycles, confirmation of each sample identification, identification of missing samples.

Laboratory: Sample reception or intake, unpacking, accessioning, time or speed to storage, temperature of storage, confirmation of sample identification, notation of any

missing samples, number of freeze–thaw cycles, processing methods, analytical environment, analytical methods, instrument(s) used, number and identity of chemists involved, batch processing status (and all related details thereof).

How do you think biobanking is improving experimental design for metabolomics experiments? Are there any disadvantages to biobanks where samples are collected with unknown purposes in mind? Is there a trade-off between flexibility and less-optimum experimental design?

Michael A. Schmidt: First, greater attention is being given to dividing specimens into aliquots in anticipation of asking a series of research questions over time. This method eliminates 1 significant contributor to experimental variance, which is freeze–thaw cycling. Also, knowledge that a sample may be stored for extended periods can encourage investigators to probe more deeply into elements that influence stability over time (e.g., whether to use −80 °C or −40 °C for storage, or another option).

Todd Mitchell: The availability of human tissue is invaluable. Some experiments, especially those examining uncommon diseases, would take decades to perform if biobanking wasn't available. There is some trade-off between flexibility and optimum experimental design. For example, analysis of some bioactive lipids requires strict collection and storage protocols, making them difficult to analyze from biobanked tissue.

Rima Kaddurah-Daouk: Use of biobanked materials has proven to provide great value including in deriving metabolomics data and biochemical insights. Metabolomics needs to adapt to realities that, within clinical settings, samples might very well have been collected without specific plans as to how they will be used. If there was a specific purpose for use on collection, then optimized collection for what is to be measured should have been attempted. Yet, there is substantial knowledge to be gained without specifications.

Tobias Pischon: State-of-the-art biobanking should reduce parts of the technical variability that are related to storage conditions of samples. In addition, regarding metabolomics, it is of course preferable to perform the analyses within a narrow time frame in a single laboratory to reduce potential effects related to batch-to-batch or between-laboratory variation. I would therefore say that state-of-the-art biobanking is an essential element to conduct metabolic experiments for any study that collects biological samples across study participants over longer periods. This particularly applies to most larger-scale studies that involve humans. For example, baseline recruitment in large epidemiologic cohort studies often takes several years. There is a likely trade-off between flexibility and specific requirements related to sample collection and storage. Thus, the better it is known a priori what metabolomic analyses the samples will be used for (and which metabolites will be analyzed), the more sample collection, processing, and storage can be tailored, which may enhance accuracy and precision. This reduces the possibility for potential analyses that go beyond the original planned design. This, however, is an issue that not only applies to metabolomic experiments per se but to studies in general.

Michael A. Schmidt: Many samples submitted to biobanks are for purposes other than metabolomics experiments and are initially focused on targeted analysis. Thus, these specimens are frequently not prepared with attention to the wide range of molecular classes and the variable stability of molecules within these classes. Moreover, contributors to experimental variance can scale substantially in the conditions of high variable and low participant numbers. Many of these potential contributors are not controlled for in studies that do not initially set out to assess the untargeted metabolome.

One of the critical trade-offs can be between broader feature diversity and the degree to which the sample best addresses the research question. For example, the metabolome derived from fecal water versus whole feces can present notably different metabolome feature profiles. Such differences should be weighed in the context of the biological question being asked. For instance, the microbial contribution to greater feature diversity in whole feces may be desirable, if this characteristic is pertinent to the research question. However, it can also be confounding if one wishes to limit the contribution of ongoing microbial activity to the analytical results. The choice of fecal water would remove the contribution of ongoing microbial metabolic activity.

Vidya Velagapudi: Currently, biobanks are improving the experimental design for metabolomics experiments, such as sample collection and storage procedures, compared to the past, when the main aim was to extract the genetic material. I think there are major disadvantages to biobanks where the samples are collected with unknown purposes. In the future, if one would like to perform, for instance, metabolomics analyses, then those samples might be useless because the analysis requires a tightly controlled procedure for sample collection and storage.

Is serum or plasma better for metabolomics studies? If using an anticoagulant, which one and why? Should we be storing and analyzing whole blood more?

Rima Kaddurah-Daouk: A substantial amount of data supports the value for profiling of both serum and plasma.

The metabolome captured in both is not identical, but similarities do exist with regard to many metabolites measured. For example, if one is interested in platelet metabolites like serotonin, use of serum is better because it captures some of what platelets produce. However, certain metabolic profiling methods are validated on EDTA or heparin plasma, and thus it is preferable if the validated biofluid is used for measurement. Whole blood collection would be extremely valuable because minimal manipulation by collectors would minimize variability.

Tobias Pischon: In my experience, most persons in the metabolomics field recommend using serum, probably because of a lower likelihood that anticoagulants affect the metabolites. As such, the answer to this question depends on what metabolites will be measured. We have done our studies mostly with serum samples, and for biobanking of samples, I would probably also recommend using serum. One should, however, not neglect the fact that because of the preparation procedures required, the use of serum may introduce additional technical variability, and therefore the preparation needs to be highly standardized.

Todd Mitchell: I have only used plasma and erythrocytes, so I don't have any experience comparing serum to plasma. In general, we used EDTA as an anticoagulant. Blood samples can usually be collected quickly and easily. In our research in which we are dealing with highly reactive compounds and if a study is specifically focused on analyzing blood, I think it is better to design your study around this and collect the samples yourself. That way you don't have to compromise, as discussed earlier. Nevertheless, biobanking blood is useful for correlating blood and tissue biomarkers when other tissues from the same donor have been collected simultaneously.

Vidya Velagapudi: In our national Metabolomics Core Facility, we have analyzed mainly serum rather than plasma samples. There is no clear-cut answer about which sample type is better for metabolomics analyses because it mainly depends on which metabolites are to be analyzed, as mentioned earlier. The most important aspect to keep in mind is to be consistent in sample collection and storage throughout the study. So far, there are relatively fewer studies on whole blood metabolomics analyses. Perhaps we should also consider analyzing whole blood more in future.

Which metabolites or lipids do you think are most vulnerable to storage changes? What is the longest length of time you think it is reasonable to store a biobanked sample and the analysis still be reliable? Which substrates do you think are most sensitive to storage changes? Should biobanks be doing anything different routinely to ameliorate this effect?

Tobias Pischon: We have done our metabolomics analyses on disease risk from samples that were collected in the 1990s and that have been stored in the vapor phase of liquid nitrogen since then. These studies worked fine, and the results were reproducible in other studies. However, our studies were not specifically designed to investigate the effect of storage time or condition on metabolite concentrations. Any effect of storage time may, of course, differ across metabolites. Therefore, I don't think that there is a specific maximum for how long such samples may be stored. However, there is very little literature on this, and studies are warranted to investigate these effects in detail.

Todd Mitchell: As mentioned above, some of the bioactive lipids, such as proinflammatory and inflammatory-resolving lipids, are highly reactive and probably need to be collected on demand.

Vidya Velagapudi: We have done polar metabolomics analyses on population cohort blood samples that were collected in the 1980s to check their quality. Samples had been stored initially at −20 °C then after some time at −70 °C. More than half of the analyzed metabolites were unstable compared to our internal quality-control samples. This implied that maintaining proper storage conditions was crucial for metabolomics analyses.

Are there any samples not being collected routinely by biobanks that you think should be? Why?

Rima Kaddurah-Daouk: Fecal material for gut microbiome studies and saliva contain substantial valuable information about the influences of microbial biomes on human health.

Tobias Pischon: There is currently a lot of interest in the potential role of the gut microbiota in health and disease. Therefore, the collection of stool samples is something that should be considered in larger studies.

Vidya Velagapudi: So far, mostly biofluid samples have been collected at biobanks owing to ease in collection and preparing the samples for storage. In my opinion, other sample types like tissues, feces, and dried blood spots should also be collected. Some disease types, such as cancer, mitochondrial disorders, and inflammatory and bowel disorders have already shown organ-specific biomarkers that play crucial roles in disease prediction and treatment outcomes. Thus, circulatory metabolites may not represent the disease activity and severity, and we may require organ tissue samples to address some biolog-

ical questions. On the other hand, dried blood spots are the best choice of sample type, especially when there is no access to the clinical laboratory or in low economic countries where there are not enough resources to easily follow protocols that require rapid processing of blood samples.

Michael A. Schmidt: Many studies of the fecal metagenome do not concurrently assess the fecal metabolome. The proper collection of fecal samples that enables both metagenome and metabolome analysis will add considerable value to our understanding of health and disease.

Todd Mitchell: The tissues I'm interested in for my research are all being collected by biobanks routinely. Sometimes patient and/or sample information are not adequate though.

There is much talk about large-scale studies. What classifies a study as large scale, and are overall number of samples or n number per group more important in this classification?

Michael A. Schmidt: Regardless of whether a large- or small-scale study, it is becoming increasingly clear that longitudinal analysis is an important means to address individual variance. Longitudinal sampling is often considered with regard to the active study period. However, there is additional value in serial baseline and end point measures, so that the baseline and end point results are not dependent upon a single point in time. For example, in the NASA Twins Study, 3 baseline and 6 end point samples were obtained (spanning 120 and 180 days, respectively).

Rima Kaddurah-Daouk: What constitutes large scale has evolved as metabolomics tools became more easily accessible and handling of large samples became possible. A decade ago we were profiling hundreds of samples and thought this was large. Currently, we are ready to profile hundreds of thousands and gearing toward millions, and that became the new large.

Todd Mitchell: As far as I know there is no set definition to determine if a study is "large scale" and is therefore somewhat subjective. Personally, I would consider any study with an "n" in the hundreds as large scale for lipidomics, although this number would be small for public health studies. The "n" per group is by far the more important factor because that is what determines the study's power.

Vidya Velagapudi: I would agree with Todd. In my experience, the "n" per group is more important than the overall number of samples in any study because the sample size is important for the downstream data analysis statistics.

Tobias Pischon: It is difficult to give an absolute number for what defines a large-scale study, and I am not sure whether we need such a definition. Although the number of samples primarily determines how many analyses can be conducted, the size of a study is a crucial design element that heavily influences the inference that can be drawn. The necessary size of a study heavily depends primarily on the study question and on the desired effect size and precision of results but also on other questions such as generalizability. It also depends on the trait of interest and its distribution in the study population. For example, discrete traits usually require a larger sample size than quantitative traits. In the field of genomics, most studies on disease risk have been conducted with case-control or cross-sectional designs. For metabolomics, however, this study design is less appropriate because it cannot differentiate between cause and effect (reverse causation bias). Thus, metabolic studies that aim to identify disease risk factors require prospective cohort studies, in which samples are being collected among participants who do not yet have the disease in question. Therefore, the required sample size in these studies depends on the incidence rate of the disease that one wants to investigate. This is the prime reason why population-based cohort studies (i.e., primary prevention settings) require much larger sample sizes than patient-based cohort studies (i.e., secondary prevention). An important reason for the large sample size required by many metabolomics studies is that they are often exploratory in nature, and therefore, multiple testing needs to be taken into account, which usually requires lower significance levels and larger sample sizes.

Are there any inherently different inclusion or collection processes that should be followed when collecting for a large-scale study or for studies that require very long collection times (e.g., rare diseases) compared to a small-scale study?

Rima Kaddurah-Daouk: To maximize knowledge gained across diseases, it is best to collect under similar conditions and with the exact same protocols. That might not be feasible. Vidya has already mentioned the drive to collect small-volume blood samples like single blood spots on paper and mailing them from across a large number of sites. This method would be ideal if technologies advance enough to make this a robust collection measure for analysis. For deeper studies on orphan diseases, it would be useful if multiple samples can be collected per specification and with optimization for each area of biochemistry being evaluated, but if not, at

least 1 sample collected per large study protocol would be helpful.

Todd Mitchell: Generally, the collection process is similar, it is just that correct storage becomes even more critical for longer term storage.

Vidya Velagapudi: In my opinion, compared to small-scale studies, when collecting biobank samples for studies that require a very long time (e.g., rare diseases), we should focus on maintaining good quality patient registries. Standardized quality assurance protocols should be followed to collect good-quality samples. Training is also a very important component here, i.e., training the biobank or clinical professionals such that all the samples would be collected in a similar fashion to avoid technical variability. Harmonization is the key point. In addition, patient or clinical registries and associated biobank sample databases should be linked correctly, and a sample information retrieval system should be present. Another crucial practical issue is to maintain sufficient funding to collect and store the samples for a very long time.

Michael A. Schmidt: I previously discussed the variables that contribute to technical variation in collecting, processing, and storing of samples in the answer to the first question. Although these variables are important in all research involving clinical chemistry, they tend to scale in large-scale metabolomics studies. Moreover, accumulating experimental variables becomes additionally problematic in metabolomics studies involving small participant numbers, such as those encountered in astronauts, athletic teams, military Special Forces, or rare diseases. Small-scale studies are particularly susceptible to overfitting owing to large variable numbers and small sample (participant) numbers. An important means to address this problem is longitudinal analysis with careful annotation of factors that contribute to experimental and biological variance. Longitudinal sampling is often considered with regard to the active study period. However, there is additional value in serial baseline and end point measures, so that the baseline and end point results are not dependent on a single point in time, which means that the experimental design of collection is important.

Although small participant numbers are never ideal, they may represent the only source of data available from which to advance certain fields. For example, in the NASA Twins Study, only 1 individual flew in space, while his monozygotic twin remained on Earth ($n = 2$). In this study, samples were collected over 27 months, including the 340 consecutive days on the International Space Station. The limiting effect of the small participant number was partially addressed by longitudinal sampling, a critical component of small "n" studies. Although not a large-scale study, this experiment has provided the most robust molecular data set to date on any human in the extreme environment of space (low Earth orbit).

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: M.A. Schmidt, Sovaris Aerospace.
Consultant or Advisory Role: None declared.
Stock Ownership: M.A. Schmidt, Sovaris Aerospace.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: R. Kaddurah-Daouk, Scientific Meetings and Scientific Reporters.
Patents: None declared.
Other Remuneration: J.A. Kirwan, University of Gottingen symposium: invited speaker, travel paid.

Previously published online at DOI: 10.1373/clinchem.2018.298620